NEW SERIES, NO. 132 (VOL. XVII) DECEMBER, 1920

QUARTERLY PUBLICATION OF THE

AMERICAN STATISTICAL ASSOCIATION

ON THE STUDY OF DISPLACEMENTS WITHIN A POPULATION

By Harald Westergaard, *University of Copenhagen*

I

Vital statistics have made rapid progress in the recent decenniums. It may be somewhat difficult for an outsider to realize this great change, but a study of the history of statistics will show it to be an indisputable fact. The progress may seem slight because the founders of political arithmetic and several generations of their disciples often anticipated results which the more scrupulous statisticians of the day have had great difficulty in proving. In fact, the modern statisticians are very careful with regard to methods as well as observations. They wish to be quite sure of the results they publish. Many of the apparent riches which the pioneers have left us and which in former days met with common assent, will be laid aside on account of their lack of real foundation. The conclusions reached by conscientious statisticians of today will, as a rule, stand the test of very severe criticism.

To characterize the present condition of statistics I shall mention two points of great importance. One is the great technical progress of which Mr. John Koren's *Collection of Memoirs of the History of Statistics*, published by the American Statistical Association, bears evidence. Undoubtedly, numerical observations nowadays are, as a rule, much more trustworthy than in former times and, moreover, we are able to specify them to a much larger extent. The second point is the constant effort to improve methods. Taking mortality statistics as an instance, it may be maintained that this discipline has reached a point of relative perfection, new methods for the construction of life tables having been devised and old methods having been improved. To a certain degree

the same will hold good with regard to birth statistics which in several countries show a remarkable evolution.*

But in the history of science it will always be observed that whenever a problem has been solved, new questions will appear on the horizon. Even in the chapter of mortality there is a multitude of unsettled questions, for instance, with regard to heredity. And in close connection with mortality statistics we meet the problems of invalidity and infirmity. Again, there is the criminal man: we need statistical investigations concerning individuals who somehow or other have had a conflict with law. Further, a most important problem is that of migration, not to speak of the transition from one class of society to another, or of changes with regard to confession or conjugal condition.

These questions will generally require fresh observations. Thus a study of heredity will require numerous data from pedigrees and other sources. But at the same time it is worth while to ask whether the old material has always been exhausted, or whether an old census report may under certain circumstances give fresh output, just as a gold mine which has been given up for years often will pay if new technical methods are applied.

Of course even the most modern census will present numerous imperfections. It is sufficient to mention the ages distribution with its maximums at the round years of age, as 20, 30, etc. But usually there will be no insurmountable difficulty in mastering these imperfections so that sufficiently safe conclusions can be drawn. Some simple process of smoothing will, for instance, often be sufficient to do away with the disturbing effect of the irregular distribution according to age.

II

In order to have a point of departure for the present investigation, let us ask whether it is possible to calculate marriage rates from the distribution of the population according to conjugal condition.

A population with its constantly varying elements may be looked upon as a function with three dimensions. At a given moment t there will be a certain number of persons aged x. If ϵ is an infinitely small element of time, then the number of persons between the ages x and $x+\epsilon$ may be representedby $y=\epsilon p_{x,t}$. The value of y will be constantly

* In one of the last numbers of the *Journal of the Royal Statistical Society*, Dr. T. H. C. Stevenson has published an investigation based on the English Census of 1911, on the "Fertility of Various Social Classes in England and Wales," giving most interesting contributions to this chapter of vital statistics. I may be pardoned for quoting an investigation with a similar object, which I published in 1890 in coöperation with Mr. Rubin, dealing with results of the Census of 1880 in Copenhagen. (*Aegteskabsstatistik*, Copenhagen 1890, translated into German under the title: *Statistik der Ehen.*) Later on the Danish Statistical Department has treated the same question in official publications, giving more recent observations for the whole country. Denmark can thus claim priority with regard to this problem.

varying on account of migrations, mortality, and births. It will, for instance, increase, say one per cent a year; its dependence on x will be still more striking, with a conspicuous decrease for each year that is added to the age. Now let the time t and the age x get the same addition ϵ; what influence will these changes have on the number of persons? The bulk of the group will evidently remain unchanged in the short interval of time, but migration and death combined may cause some small change, generally a decrease. Let δ_x be the rate of decrease from these causes, corresponding to the force of mortality; then we shall have:

$$p_{x+\epsilon,\, t+\epsilon} = p_{x,\, t}(1 - \epsilon\delta_x),$$

or

$$\frac{dp_{x,\,t}}{dx} + \frac{dp_{x,\,t}}{dt} = -p_{x,t}\,\delta_x.$$

Of course δ_x is dependent not only on the age, x, but also on the time, t, though in many cases we may be entitled to look upon this quantity as constant in the interval between two censuses.

Let, further, $\epsilon b_{x,\,t}$ be the number of bachelors aged x to $x+\epsilon$; then supposing their rate of migration and death to be δ'_x and the rate of marriage to be m_x, we shall have:

$$b_{x+\epsilon,\, t+\epsilon} = b_{x,\,t}(1 - (\delta'_x + m_x)\epsilon),$$

or

$$\frac{db_{x,\,t}}{dx} + \frac{db_{x,\,t}}{dt} = -b_{x,\,t}(\delta'_x + m_x).$$

Finally, we can put $f_{x,\,t} = \dfrac{b_{x,\,t}}{p_{x,\,t}}$, $f_{x,\,t}$ being the ratio between the number of bachelors and the total population. We get the following equation:

$$\frac{df_{x,\,t}}{dx} + \frac{df_{x,\,t}}{dt} = -f_{x,\,t}(\delta'_x - \delta_x + m_x).$$

As a rule it will not be difficult to calculate the expression on the left side of the equation with sufficient approximation. On the right side of the equation we have three unknown quantities, and the value δ_x is actually a sum of two quantities, say i_x the rate of migration and μ_x the force of mortality and the like, then we shall have $\delta'_x = i'_x + \mu'_x$. There are consequently as many as five quantities, and it might therefore be hopeless to try to find the marriage rate for bachelors. Nevertheless, in this case we are not prevented from reaching an approximate result. Usually bachelors will have a greater rate of mortality than married people, and the rate of migration will probably be greater also.

We shall thus very often have $\delta'_x > \delta_x$, and leaving these quantities out of consideration, we get a value for m_x, which as a rule is too high. But the difference between this value and the real one will generally be small. In the period of life in which most marriages take place, the rate of marriage will be 10 per cent or more, whereas the force of mortality and probably often the rate of migration, also, will be only a few per mille. If this is the case, the difference $\delta'_x - \delta_x$ cannot as a rule count much, and we will therefore often be justified in putting as a first approximation:

$$m_x = -\frac{1}{f_{x,t}}\left(\frac{df_{x,t}}{dx} + \frac{df_{x,t}}{dt}\right).$$

That is to say, if we know only the relative number of bachelors according to the census reports, we are able to find their marriage rate without knowing the number of marriages taking place.

In many cases it will even be justifiable to simplify the above equation still more by omitting $\frac{df_{x,t}}{dt}$. This may be necessary, if there is only one census available, or if the observations in the various census reports are not homogeneous. To see the extent of these changes in modern times we may look at the following table for Denmark, 1901 and 1911, giving the proportionate number (per 1,000) of bachelors and spinsters at different ages:

	Males		Females	
Age	*1901*	*1911*	*1901*	*1911*
20–25	881	876	746	727
25–30	497	488	418	400
30–35	246	235	252	254
35–40	151	145	191	194
40–45	106	115	156	173
45–50	89	96	135	155
50–55	77	85	127	138

If these numbers were constant from one census to another, they would tell us how a generation of bachelors or spinsters, if they were not exposed to mortality or migration, would by and by be reduced by marriage, and one census would be sufficient to calculate the approximate values of the rates of marriage. Evidently there is, however, a movement testifying to an increasing frequency of early marriages and a decreasing frequency of marriage in advanced years. From the numbers above we should find for males 20–25 years old with the year as unity:

$$-\frac{1}{f_{x,t}} \cdot \frac{df_{x,t}}{dt} = 0.0006.$$

In the following quinquennium of age we get 0.0018; the values are larger in the more advanced years, and on the whole more considerable for women than for men.

Still, these quantities are small compared to the variations with regard to the age x. The decrease from one quinquennium of age to another is so great that it will be necessary to take some care in the calculations. It may be advisable to go back to the values of $b_{x,t}$ and $p_{x,t}$, calculating their variations separately. As to the unit, it will generally suffice to calculate the relative numbers of bachelors for each year of age, and to use these numbers as approximate expression for $f_{x,t}$ for the median age, for instance 22, 5, 23, 5, etc. It will be advisable to let the numbers undergo some light smoothing process before calculating the rates of marriage. We thus reach the following results:

	Calculated Rates 1911		Observed Rates 1911–15	
Age	*Males*	*Females*	*Males*	*Females*
17.5	....	0.019	....	0.016
22.5	0.069	0.110	0.063	0.103
27.5	0.149	0.113	0.156	0.119
32.5	0.135	0.078	0.127	0.063
37.5	0.073	0.029	0.080	0.033
42.5	0.032	0.019	0.041	0.017
47.5	0.019	.....	0.024	0.009
52.5	0.020	0.010	0.012	0.004

In comparing the results of the calculation with the observed numbers, we must bear in mind that the two columns do not cover exactly the same ground, one of them giving the rates at the moment of the census only (1 Feb. 1911), the other one giving the average 1911–15. Moreover, the observed rates are calculated for the ages 20–25, 25–30, etc., not for single years of age: finally, the various stages of calculation (including the variations with regard to t), as well as the whole smoothing process, may give rise to some difference. We must remember, as remarked above, that there is reason to consider the rates that have been found in this way as somewhat too high, the difference $\delta'_x - \delta_x$ having been left out of consideration. If we calculate the number of marriages which would take place in a population like that of 1911, according to the calculated and the observed rates, the former will give 4 per cent more marriages among men than the latter, and about 5 per cent more marriages among women. If we had no knowledge from direct observation, of the rate of marriage of bachelors, we might be justified in looking upon these calculated rates as a rather fair approximation. In some countries rates of marriage for the whole population are known, but not for widowers and bachelors separately. The

described method would then help us to draw conclusions with regard to the rates of marriage for widowers as well as for bachelors.

It might be objected that all these calculations are superfluous if we are possessed of satisfactory direct observations, and that we have only to collect such observations in order to find the rates we wish to know. But this cannot always be the case. For example, we are in possession of a very interesting old Danish census of 1787, but we have no corresponding knowledge of the marriages, so that if we wish to form a picture of the chances of marriage of those days we shall have to resort to devices of the described character.

This census of 1787 is in many respects a most remarkable document; it gives—with 10 year intervals of age—the combined distribution of population according to sex, conjugal condition, and occupation. The numbers reveal the most striking difference between the old society and the modern one. Thus in the age of 30–40 among 100 males, no less than 29 were bachelors in 1787 as against 19 in 1911. In the following 10 year group there was a difference in the opposite direction, 9 against 12, and among females the difference was still greater, the corresponding numbers being 9 and 16.

These enormous changes are, however, chiefly a modern product. In 1801 the relative number in the age of 20–30 was 80 for males, the same as in 1787; for females the number was reduced from 64 to 63. In the following age-group the numbers were unchanged. Thus, in order to find the rates of marriage it will be justifiable to leave the variations with regard to time out of consideration.

The rates of marriage have been found for the median ages 25, 35, etc., by supposing the number of bachelors as well as the total number of persons to be represented by algebraic functions. Putting $b_x = m + nx + px^2$, we find, the central age being 0,

$$a = \int_{-15}^{-5} (m+nx+px^2)dx, \quad b = \int_{-5}^{+5} (m+nx+px^2)dx$$

and

$$c = \int_{+5}^{+15} (m+nx+px^2)dx$$

where a, b, c, are three consecutive observed numbers of bachelors, and it is then easy to find m, n, and p; the differential coefficient of b_x for $x=0$ will be n. In the same way we can deal with the total numbers A, B, and C, and we find as an approximate value of the marriage rate:

$$\frac{6}{5}\left(\frac{C-A}{26B-A-C} - \frac{c-a}{26b-a-c}\right);$$

or we may choose a still simpler formula on the supposition that we can put $2b=a+c$ and $2B=A+C$: the formula will then be reduced to

$$\frac{C-A}{20B}-\frac{c-a}{20b}.$$

In this way the following approximate values of the rates of marriage per mille in 1787 have been found:

Age	*Males*	*Females*
25	39	53
35	138	185
45	132	92

The numbers show a striking contrast to the rates for 1911–15, given above. In the younger years there were proportionally few marriages, but after 30 the opposite is true.

To understand these enormous differences it will suffice to consider certain observations in various classes of society. For instance, we find among craftsmen in the town population the following absolute numbers:

	Masters		Journeymen	
Age	*Married and Widowers*	*Bachelors*	*Married and Widowers*	*Bachelors*
20–30	511	302	321	3,578
30–40	1,748	173	674	876
40–50	1,589	88	397	187
50–60	1,153	56	233	68

Taking masters and journeymen together we find that in the age of 20–30 years about 82 per cent were unmarried; in the following age group the number is only 30 per cent, between 40 and 50 it has been reduced to 12, and between 50 and 60 to 8. The married people were to be found chiefly among the masters. Thus, whereas 57 per cent of the journeymen between 30 and 40 were bachelors, the corresponding number among masters was only 9 per cent. It appears, further, from these observations that there was a constant stream from the class of journeymen to that of masters, but many journeymen reached the position of independent householders very late in life, and consequently marriage had to be put off. In most professions the journeyman lived in his employer's house, as a member of his household, till he settled as master. It is quite otherwise in modern society; most workingmen have no chance of becoming employers, so that they have very little inducement to put off marriage. It may even be the case that a young workingman aged 20 or 25 is frequently better off than he is in more advanced years.

III

Returning to the equation:

$$\frac{df_{x,t}}{dx}+\frac{df_{x,t}}{dt}=-f_{x,t}(\delta'_x-\delta_x+m_x),$$

we can use $f_{x,t}$ to signify other groups within the population. In the parenthesis, m_x will be the frequency of cases which tend to diminish the group, being negative, if there is an increase of the number. As with the marriages there will sometimes be the simplification that the variation with regard to t can be considered as zero, so that $\frac{df_{x,t}}{dt}$ can be left out of consideration. As an instance we can choose the feeble-minded. Under this heading the Danish census chiefly registers persons who are feeble-minded from birth or from a very early age. It is a mere exception that persons registered as feeble-minded are married. If the feeble-mindedness had often appeared later in life, the number of married persons in this group would undoubtedly be much larger. Thus m_x will probably be zero or at least very low (being an increase to the group, it will not be positive as the marriage rate). The variations of $f_{x,t}$ with regard to t will probably also be very small, though the census reports seem to indicate the reverse; the explanation seems to be that previous census reports have been less successful in the enumeration of the feeble-minded than was the census of 1911. Probably migration plays only a trifling part among idiots. In the parenthesis therefore we have chiefly to consider $\mu'_x-\mu_x-i_x$, μ'_x being the force of mortality among feeble-minded and μ_x among all persons, i_x the rate of migration. If fresh cases have a noteworthy effect, we shall have to add the negative quantity m_x. In the following table are shown the numbers of feeble-minded in Denmark, observed in 1911, among 10,000 of each age and sex, the numbers being adjusted by taking the average of three values as expression for the median:

FEEBLE-MINDEDNESS IN DENMARK, 1911

Age	*Males*	*Females*
15–20	29.3	25.3
20–25	28.1	22.7
25–30	26.1	20.5
30–35	24.0	19.7
35–40	21.4	17.8
40–45	19.2	16.4
45–50	17.0	13.1
50–55	14.7	11.4
55–60	13.1	10.3
60–65	11.9	9.4
65–70	9.3	8.0

The yearly decrement in the age 20–25 would be about $1/10(29.3-26.1)=0.32$. Dividing by 28.1 we find a quotient 0.011. The normal rate of mortality being 0.004, the mortality of feeble-minded will be at least 0.015. In this way the following table has been calculated:

Rate of Mortality Per 1,000

	Idiots		Normal Persons	
Age	*Males*	*Females*	*Males*	*Females*
20–25	15	25	4	4
25–30	20	19	4	4
30–35	25	19	5	5
35–40	28	25	6	6
40–45	30	35	7	6
45–50	36	46	10	8
50–55	40	35	13	10
55–60	40	33	19	14
60–65	59	45	27	21

These numbers can of course give only very rough estimates, but we can hardly read them otherwise than as testifying to a very high mortality among idiots. I may here refer to an investigation which I made several years ago concerning the mortality of idiots in asylums for feeble-minded, thus probably singling out a group still more delicate in health.* I found a mortality which was on an average nearly double the rates stated above; this seems to corroborate the results from the method here explained.

Turning to the insane, we find particulars of the same kind in the Danish census reports. In the following table are given the unadjusted numbers per 10,000 according to the census of 1911:

Insane Persons

Age	*Males*	*Females*
15–20	1.9	2.3
20–25	9.4	9.3
25–30	20.6	16.6
30–35	27.7	22.4
35–40	32.8	35.6
40–45	40.3	40.7
45–50	38.7	53.5
50–55	43.8	53.1
55–60	44.0	57.4
60–65	46.7	51.8
65–70	48.9	58.5
70–75	46.1	47.3

It follows from these numbers that in a long period of life there is a constant increase of lunatics, in spite of the recoveries and in spite of

* *Die Lehre von der Mortalität und Morbilität*, 2te ed., 1901, p. 237.

their undoubtedly high mortality. Thus between 20 and 25 we find for males a yearly net increase of about 20 per cent of the existing number. If we prefer to ask what is the probability of a young man's becoming insane at the age 20–25, we shall have to compare these fresh cases to the number of non-lunatic men, viz., 9,991, and the lower limit of the probability of becoming insane within a year will be 0.2 per mille. But in this case the excess of mortality and the frequency of recovery will play such an important part that it is less satisfying to follow the described line of thought with regard to this material.

With regard to the blind and the deaf and dumb, the Danish census of 1911 gives the following numbers per 10,000:

	Blind		Deaf and Dumb	
Age	*Males*	*Females*	*Males*	*Females*
Under 5 years	1.2	1.1	1.4	1.3
5–10	1.9	1.8	4.8	3.6
10–15	3.5	3.0	7.2	8.0
15–20	3.1	2.5	9.0	5.7
20–25	3.2	2.0	7.4	6.2
25–30	2.5	2.8	9.8	7.0
30–35	3.6	3.0	8.7	6.1
35–40	3.2	4.0	14.5	9.3
40–45	4.1	4.7	8.3	6.8
45–50	6.1	4.8	7.5	6.4
50–55	8.3	6.7	8.3	5.0
55–60	10.3	8.0	7.1	7.3
60–65	12.2	9.6	9.0	6.6
65–70	16.2	17.2	9.2	7.7
70–75	20.3	22.4	5.9	5.9

The numbers of deaf and dumb depend to some extent on epidemics. Thus in Denmark the abnormal numbers between 35–40 in the year 1911 can partially be ascribed to epidemics from meningitis cerebrospinalis in Jutland, 1872–73.* Leaving out this group we find on inspection of the relative numbers that the frequency of this defect in the ages 15–35 is about the same as between 40 and 70. On the whole there seems to be no striking difference between the mortality of the deaf and dumb and of the general population. This was confirmed by a direct investigation by Dr. Mygind several years ago.†

The numbers concerning the blind are more complicated. The function $f_{x,t}$ cannot be looked upon as constant with regard to t, for there has been a conspicuous decrease in the blindness of newly-born owing to hygienic and medical progress. Thus the relative numbers in advanced years would be larger than if the present chances of the blind were used. In the equation

* Westergaard, *loc, cit.*, p. 239.

† *Nationalökonomisk Tidsskrift*, 29, vol. 1891.

$$\frac{df_{x,t}}{dx}+\frac{df_{x,t}}{dt}=-f_{x,t}(\mu'_x+i'_x-\mu_x-i_x+m_x)$$

we may be entitled to leave $\mu'_x+i'_x-\mu_x-i_x$ out of consideration. The negative quantity m_x will thus mainly depend on the chances of becoming blind, as there is a very small chance for a blind person to recover; if we take no notice of the variations with regard to t, we shall somewhat exaggerate the probability. After 15 the frequency of blindness is nearly constant for a long series of years, but at about the age of 40 an increase is beginning. We find, for instance, that out of 100,000 the following numbers would yearly become blind:

Age	*Males*	*Females*
55–60	4	8
60–65	6	10
65–70	8	17

There is, therefore, an increasing probability with increasing age, and the female sex appears much more liable to blindness than the male.

It is unfortunately not possible to find the variations with regard to t, as the previous census report is not sufficiently complete with regard to the cases of blindness. But there is reason to believe that this imperfection of the material has comparatively slight influence on the numbers concerning the more advanced years of age which have been treated here.

Incomplete as all these observations on infirmity and invalidity are, they may serve as examples of the method explained, showing various possibilities under the discussion of the equation rendering the variations of the relative numbers.

IV

When two consecutive census reports are fairly correct, it may often be sufficient, instead of calculating the probabilities, merely to study the variations in the relative numbers. This will hold good with regard to the Danish religious statistics. The census registers the number of adherents to each church. The bulk of population is registered as belonging to the established (Lutheran) church, only small minorities belonging to other confessions. There is reason to believe that these minor groups have very much the same mortality as the total population, and that also the migrations with few exceptions (viz., the Jews and possibly also the Roman Catholics) are mostly of the same character. We are thus entitled to expect that the value of f_x in 1901 will be the same as that of f_{x+10} in 1911, if there are no particular movements influencing the numbers. Thus among Baptists we find in 1901, 455 males registered, aged 35–50 years. Ten years later there

were 406, aged 45–60. In the whole population the corresponding numbers were 189,891 and 169,512. Out of 1,000 persons in 1901 there were 893 left in 1911. Calculating 993 per mille of the observed number of Baptists in 1901, we find 406, which is exactly the number observed in 1911. It thus seems probable that very little change has taken place in Denmark in this church in the ages concerned. In the following, I shall give some observations of this kind:

REFORMED CHURCH

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	99	89	15–25	89	88	104	84
15–20	55	47	25–30	47	46	61	44
20–35	115	122	30–45	113	108	122	114
35–50	98	125	45–60	85	87	102	113
50–80	131	131	60–70	66	82	83	87
	498	514		400	411	472	442

The difference between the observed and expected numbers is not very striking. The Reformed Church has, on the whole, kept its members.

The Methodists also show rather small changes, as will appear from the following numbers:

METHODISTS

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	468	482	15–25	270	415	427	457
15–20	125	178	25–30	106	104	170	166
20–35	258	443	30–45	280	240	449	414
35–50	329	451	45–60	277	293	430	409
50–80	254	392	60–70	210	159	293	260
	1,434	1,946		1,143	1,211	1,769	1,706

There seems to be some decrease in younger years, whereas in riper age the numbers are somewhat increasing.

Also, the Baptists show some stability between the two census years; in younger years, however, some decrease is perceptible.

BAPTISTS

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	575	647	15–25	349	509	479	613
15–20	178	235	25–30	136	148	195	219
20–35	419	615	30–45	403	391	605	574
35–50	455	564	45–60	406	406	537	512
50–80	479	654	60–70	323	300	396	434
	2,106	2,715		1,617	1,754	2,212	2,352

Compared to these results several other churches show great variations, for instance, the Catholic-Apostolic Church (the "Irvingites").

"IRVINGITES"

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	479	486	15–25	201	424	254	461
15–20	154	193	25–30	70	128	115	180
20–35	293	429	30–45	193	273	268	401
35–50	317	392	45–60	193	283	272	356
50–80	350	318	60–90	149	157	171	211
	1,493	1,818		806	1,265	1,080	1,609

The loss is perceptible in all periods of age, but especially in younger years. On the other hand, the Roman Catholic Church has made rapid progress, as the following numbers will prove:

ROMAN CATHOLIC CHURCH

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	552	623	15–25	1,239	489	1,265	591
15–20	228	242	25–30	469	189	516	225
20–35	591	769	30–45	890	551	1,094	718
35–50	488	630	45–60	465	435	624	571
50–80	282	387	60–90	193	177	272	257
	2,141	2,651		3,256	1,841	3,771	2,362

Also here we see that the increase takes place principally in the younger years. Part of the increase is probably caused by Polish agricultural laborers coming to Denmark for a shorter or longer time.

The Jews show interesting phenomena. The small mosaic congregation in Denmark seemed for a couple of generations to be in stagnation, or even to be gradually dying out. But of late years there has been an afflux from abroad, especially from Poland and Russia. This accounts for the peculiar changes in the age distribution which can be observed. In 1901 only 18 per cent of the total number were under 15, but in 1911 the proportion had risen to 26 per cent. In the same way as above the following results have been found:

JEWISH CONFESSION

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	215	220	15–25	379	190	448	209
15–20	124	115	25–30	253	103	241	107
20–35	364	411	30–45	512	339	533	384
35–50	389	408	45–60	362	347	416	370
50–80	412	516	60–90	239	258	351	342
	1,504	1,670		1,745	1,237	1,989	1,412

The increase is very conspicuous in the younger years; in more advanced ages foreign Jews have evidently not settled in Denmark to any extent in the period which we have here before us.

Finally, I shall give the main results with regard to persons who have been registered as not belonging to any confession. The number has been rapidly increasing, though it is still very small compared to the total population. The motive for declaring withdrawal from the established church or other confessions may of course lie deep within the personality; but very often some outward event will be taken as occasion, for instance, the wish to enter into civil marriage. The total number increased from 3,628 to 8,151, in 1911.

Belonging to No Church

1901			1911				
Age	*Males*	*Females*	*Age*	*Males*		*Females*	
				Observed	*Expected*	*Observed*	*Expected*
5–15	231	178	15–25	757	205	353	169
15–20	119	63	25–30	1,149	99	270	59
20–35	1,087	327	30–45	2,347	1,013	566	305
35–50	778	212	45–60	709	694	228	192
50–80	251	97	60–90	169	154	67	64
	2,466	877		5,131	2,165	1,484	789

After the age of 45 there is no great change, whereas in younger years the increase is appreciable, especially among males aged 25–30, the number rising from 99 to 1,149.

It is not difficult to find approximate net values of the rates of withdrawal from the church. Let $b_{x,t}$ be the number of persons outside the church, and $p_{x,t}$ the number belonging to some confession, and let the rate of withdrawals over and above the rate of reëntrance be $w_{x,t}$, then in a given element of time ϵ, the number of withdrawals will be $w_{x,t}p_{x,t}\epsilon$. If a census has been taken at 10-year intervals, the first census giving $p_{x,t}$, the latter $p_{x+10,t+10}$, and if the time passed is $t+n$, then we shall have as the expected number of withdrawals: $w_{x+n,t+n}$ $p_{x+n,t+n}\epsilon$. Of these persons some will disappear in the period preceding the next census. It may be supposed that $p_{x+n,t+n}$ will be reduced to $p_{x+10,t+10}$ and the total gain found at this moment will be

$$p_{x+10,t+10}\int_0^{10} w_{x+n,t+n}dn,$$

or approximately

$$10 \cdot p_{x+10,t+10} \cdot w_{x+5,t+5}.$$

As an example we can take the gain in 1911 for males aged 25–30, altogether 1,050 or 105 per year between 1901 and 1911. The rate of withdrawals for the ages 20–25 will then be approximately 0.0011, the number of males aged 25–30 being in 1911 about 99,000, of whom

98,000 belonged to some confession. At the age of 15 we shall find 0.0002; for 30–35 years we shall have 0.0006. Above 40 the rate of withdrawals is reduced to a very small value, the excess being, according to the table, only 0.00001 per year.

V

Returning to the old Danish census of 1787, we may ask what displacement according to this census took place from one class of society to another. Practically, as remarked above, we may at this epoch look upon the population as being stable, the same proportions remaining from one year to another, even though there has been a small yearly natural increase of the population. Let $f_{x,t}$ be the relative proportion of some class to the whole population, then we may suppose that the differential coefficient $\frac{df_{x,t}}{dt}$ is approximately zero. Thus, the rural population had nearly the same proportion to the town population as 14 years later. By the census of 1901 the rural districts had, as in 1787, a population about four times the urban one, the urban population, as in the early census, being equally distributed between the capital and the provincial towns.

We may now ascertain to what extent the attraction of the towns has acted upon the population by using a similar reasoning as above with regard to the confession. For 1,000 males in the whole country aged 20–30 we should find 859 aged 30–40. Consequently the number of males in Copenhagen aged 20–30 would be reduced from 10,952 to 9,408 in the following decade, if there were no peculiar causes influencing the numbers. The actual number being 8,792, there is a difference of 616, probably owing to migrations in the age between 20–30 to 30–40, say about 30 years of age. For 100 persons expected, there were actually found only 93. In this way the following table has been calculated, showing how many persons were observed for 100 expected:

	Males			Females		
Age	*Copenhagen*	*Provincial Towns*	*Rural Districts*	*Copenhagen*	*Provincial Towns*	*Rural Districts*
10–20	107	98	99	96	96	101
20–30	170	122	90	152	129	92
30–40	93	96	102	100	96	101
40–50	78	92	105	85	98	103
50–60	82	95	103	90	102	101
60–70	71	89	104	88	102	101
Above 70	89	100	101	88	107	100

Thus, town life had a considerable attraction on males in younger years, more than counterbalancing the higher mortality which then as a

rule menaced the life in towns. But soon after the tide turns and, as the combined effect of the greater mortality and of the migrations, the numbers are smaller in advanced years of age in Copenhagen and the provincial towns than in the country. At about 20 the male Copenhagen population had an increase of about 4 per cent per year from migration, whereas the rural districts were losing only 1 per cent. Also, the female population was attracted in younger years by the towns, the reaction being perceptible in the age of 30–40.

Still more interesting were the changes which took place within the artisan class. Only a few professions were allowed in the country, such as weavers, blacksmiths and wheelwrights, the towns having the privilege of binding most professions to them. The village artisans in the census report are given in total, whereas the artisans in towns are divided into two classes, the masters (with their male children), and the journeymen and apprentices.

The following table gives the main observations in absolute numbers for males:

	Towns			Rural Districts
Age	*Masters*	*Journeymen*	*Total*	*Total*
0–10	3,391	951	4,342	6,988
10–20	1,119	3,060	4,179	2,603
20–30	813	3,899	4,712	3,232
30–40	1,921	1,550	3,471	4,184
40–50	1,677	584	2,261	3,449
50–60	1,209	301	1,510	2,748
60–70	640	144	784	1,543
70–80	188	40	228	472
80–90	26	9	35	69
90–100	2	....	2	....
	10,986	10,538	21,524	25,288

It will be seen on inspection of this table that the ranks of the village artisans are very much reduced in a certain period of life; among 10 and 20, probably a large number of this group gained their living as servants, later returning to their profession. Some village artisans have possibly spent some time in the towns.

We can now try the same calculation as above for the towns and the rural districts, and we thus arrive at the following results, showing what percentage of the expected numbers was actually found in the various groups of age:

	Towns			Villages	Total of Artisans
Age	*Masters*	*Journeymen*	*All Artisans*		
10–30	54	181	123	76	99
30–50	159	39	84	124	103
50–100	90	75	86	96	92

As observed above, the country has a loss in the first period of life and a gain in the later period. On the other hand the towns gain in the younger years but lose afterwards. From material so defective as this it will, of course, be impossible to trace all the causes. Part of the loss in the higher ages in towns may be due to the unfavorable mortality; also, a number of former artisans may possibly be found in poorhouses.

We might have tried other methods of comparison. For example, it would be tempting to calculate the numbers which would be found if all the classes in view had the same age-distribution, taking as basis the number in each class pertaining to the youngest group of age. But the picture seems to be clearer if we choose the comparison which has been explained here.

The numbers of masters and journeymen show interesting features, as was remarked above, the stream going from the journeymen to the masters. For the sake of brevity I shall, however, omit entering into more details with regard to this question. But some observations concerning the structure of the rural population may prove instructive.

This structure of village society was at that time very simple. The majority of the population (altogether 330,000 males) were peasants (tenants and freeholders) with cottiers, servants, and laborers. About 135,000 males of all ages were registered as peasants with their children, holding lands of various size, sufficient to support a family; 35,000 were small-holders and 6,000 were registered as cottiers without land, working for the peasants or in the manors. There were 26,000 laborers, some of whom probably were engaged in work other than agricultural; this has been the case, too, with some of the 67,000 male servants. About 7,000 male persons were registered as poor. The following table will show the distribution according to ages:

Males in Rural Districts 1787

Age	*Peasants*	*Small-Holders*	*Cottiers*	*Servants*	*Laborers*	*Poor*	*The Six Classes Combined*
0–10	38,936	11,059	2,142	2,479	8,250	1,822	64,688
10–20	22,545	2,229	408	25,683	1,717	501	53,083
20–30	15,089	2,011	461	25,380	1,417	176	44,534
30–40	17,869	5,163	741	8,355	4,005	233	36,366
40–50	16,863	5,613	556	2,867	4,407	349	30,655
50–60	12,347	4,684	599	1,299	3,691	697	23,317
60–70	7,767	3,074	779	466	2,046	1,442	15,574
70–80	3,106	1,157	428	94	432	1,458	6,675
80–90	491	154	84	7	21	393	1,150
90–100	16	5	7	....	2	39	69
	135,029	35,149	6,205	66,630	25,988	7,110	276,111

It will easily be seen, on inspection of this table, how people floated from one group to another, for instance, swelling the class of servants between age 10 and 30 and diminishing the number of peasants, and still more that of laborers. If we calculate the expected numbers according to the age-distribution in the whole country, we obtain the following results:

Expected Numbers

Age	*Peasants*	*Small-Holders*	*Cottiers*	*Servants*	*Laborers*	*Poor*	*The Six Classes Combined*
10–20	29,046	8,250	1,598	1,849	6,155	1,359	48,257
20–30	21,553	2,130	390	24,553	1,641	479	50,746
30–40	12,961	1,727	396	21,801	1,217	151	38,253
40–50	14,438	4,172	599	6,751	3,236	188	29,384
50–60	12,546	4,176	414	2,133	3,279	260	22,808
60–70	7,766	2,946	377	817	2,322	438	14,666
Above 70	3,726	1,259	396	205	907	863	7,556
	102,036	24,660	4,170	58,109	18,757	3,738	211,670

In the age of 10–20, the peasants have lost about one-fourth of their number, but the small-holders have lost nearly three-fourths, being reduced from 8,250 as expected, to 2,229 as observed. This is also the case with the landless working men (the cottiers and the laborers); also from the poor there is a constant stream, probably chiefly going to the class of servants. The five classes here mentioned have lost about 19,000, whereas the class of servants has gained nearly 24,000. Probably the ranks of the servants were also recruited from other classes, as that of the artisans; even the town population may have contributed a little.

In the following decennial age group (20–30), a loss is perceptible if all the six classes are taken together, whereas the class of servants has still a small increase. There is a considerable decrease in the peasant class, probably partly owing to the military service, the class of soldiers and sailors in this period of life showing a large addition.

In the next period of life (30–40) tide is turning. Five thousand persons are streaming back to the peasant class, proportionally still more to the small-holders and laborers. Altogether, the class of servants has lost 13,000 or nearly two-thirds of its expected number. There is still a small loss in the six classes, which, however, in the following period of life changes into a small increase.

The age of 40–50 shows in the main the same features as the foregoing one—the servants losing more than half of their number, whereas the other classes are winning. The decrease in the servant class is still perceptible in the following age-groups, but now this class has been so much reduced that its movements can have hardly any influence on

the other classes; thus the number of peasants seems mainly constant. A considerable increase is shown in the poor class. The following abstract will give the main results. It tells what percentage of the expected number was actually found:

Age	*Peasants*	*Small-Holders*	*Cottiers*	*Servants*	*Laborers*	*Poor*	*The Six Classes Combined*
10–30	74	41	44	193	40	37	99
30–50	127	183	130	39	189	172	99
50–100	99	106	160	59	95	258	104

In order not to make this description too fatiguing, I shall confine myself to a corresponding abstract for the female population:

Age	*Peasants*	*Small-Holders*	*Cottiers*	*Servants*	*Laborers*	*Poor*	*The Six Classes Combined*
10–30	78	58	71	187	53	56	100
30–50	114	157	143	26	160	208	96
50–100	96	94	123	53	80	183	103

On the whole we have here the same features as for the male population.

The method used here to follow the movements of the various classes of population can, under certain circumstances, be applied to modern census reports. The well-known English *Supplementary Reports on Births, Deaths, and Marriages* give good information concerning the varying ages distribution in a large number of professions, and it will not be without value to look upon these observations from the point of view here described. As another instance I shall mention an investigation which I undertook many years ago in coöperation with Mr. Rubin; it gave us the opportunity of entering into the same question.*

In order to find the rates of mortality in various classes of the rural population, we undertook a study of the original schedules of the census of 1880 for one of the Danish provinces so that we could ascertain the distribution according to ages of the classes which we wished to investigate. But whereas these observations might be looked upon as fairly correct they were not truly representative without a correction for the corresponding number of deaths. There was reason to believe that several young servants, if attacked by a serious disease, as tuberculosis, left their service and returned to their homes to die there, and if the certificate of death then did not mention the deceased as a servant, the consequence would be that the mortality of the class concerned was overrated. In order to see the extent of this source of errors, it was therefore necessary to correct the numbers as far as possible by

***Landbefolkningens Dödelighed.* Copenhagen, 1886.

supplementing the numbers of living persons and of deaths in each class of society with a probable number originating from the class of servants. Thus, among artisans the mortality between 10 and 15 would be reduced from 3.8 per mille to 3.2 per mille and between 15 and 20 from 4.8 per mille to 4.0 per mille. The ages distribution had much resemblance to the table for 1787, though probably the numbers for 1880 are somewhat less stable.

As to the census report for 1787 it would be easy to multiply applications of the method here explained. Thus, sailors and fishermen form one interesting group, the soldiers another one; but what I have developed here will probably prove sufficient to form an opinion concerning the method.

It must be acknowledged that several of the calculations described above have a less firm character than we should wish for in modern statistics. However, they will have this advantage over the old political arithmetic with its often very fantastic results, that they are supported by observations abounding in thousands of volumes of official statistics all over the world. From this stronghold it may be allowed to make excursions in fields which have not yet come under the plough of regular statistics.

VI

A little nearer the shelter of official statistics we meet the problem of calculating the rates of migration of a population, the rates of mortality having been fixed with tolerable accuracy from the lists of deaths and the census reports. In fact, knowing as we do the number of deaths of a certain generation between two moments of census, we can easily calculate the gain or loss from migration, and thus we have material for calculating the rates of migration according to age.

If the force of mortality at a given age x is μ_x and the rate of the net-emigration is i_x, then we have:

$$\delta_x = \mu_x + i_x$$

where δ_x is found by the equation:

$$p_{x+\epsilon,\, t+\epsilon} - p_{x,\, t} = -\delta_x p_{x,\, t},$$

knowing μ_x, we get i_x by subtraction.

From this equation we get the following:

$$p_{x+10,\, t+10} = p_{x,\, t} \cdot e^{-\int_0^{10} \delta_{x+n} dn},$$

or approximately

$$e^{-10\, \delta_{x+5}} = \frac{p_{x+10,\, t+10}}{p_{x,\, t}}.$$

The main results for Denmark for 1900–10 will be the following for males:

Age	*Force of Mortality Per Mille*	*Rate of Net-Emigration Per Mille*
5–10	3	3
10–15	2	5
15–20	3	15
20–25	4	15
25–30	5	4
30–35	5	1

After 45 the rate of migration is very small. It will be seen from these numbers that it is principally the young men who leave their country.

As to the female population between 5 and 25, the yearly excess of emigration over immigration will be about 3 per mille; after this age for a long period of life the rate is reduced to 1 or 2 per mille. The influence of age on migration is thus much less than among males.